DOCUMENT RESUME

ED 408 300 TM 026 501

AUTHOR Guskey, Thomas R.

TITLE The Relationship between Socioeconomic Characteristics and School-Level Performance Assessment Results.

PUB DATE Mar 97

NOTE 30p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 1997).

PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS *Academic Achievement; Achievement Tests; *Educational Assessment; Elementary Secondary Education; Evaluation Methods; Low Income Groups; *Performance Based Assessment; Portfolios (Background Materials); Poverty; Regression (Statistics); Scores; *Socioeconomic Status; State Programs; *Student Characteristics; Test Results; Testing Programs

IDENTIFIERS *Kentucky Educational Assessment Program

ABSTRACT

The relationship between selected contextual and socioeconomic variables and school-level results from a statewide performance-based student assessment system was studied using the Kentucky Instructional Results Information System (KIRIS). KIRIS uses portfolios of student work, scores from on-demand assessments, and student performance on performance events in the areas of mathematics, science, social studies, arts and humanities, practical living, and vocational skills. Multiple regression analyses were conducted using data gathered over a 3-year period (1993-1995) from 49 schools in one school district. Results show that a single socioeconomic variable, the percent of students qualifying for free or reduced lunch benefits, explains a large portion of the variance in scores at all school levels. Contextual or socioeconomic indicators were not, however, predictive of improvements in scores from year to year. Although socioeconomic indicators explain much of the variation in elementary, middle, and high schools' levels of achievement, they do not explain the variation in score improvements made by schools in this district. Possible explanations for these results and school level differences are discussed, along with implications for performance-based student assessment systems. (Contains 5 tables and 21 reference.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 408 300

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
THOMAS R. GUSKEY
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

The Relationship Between Socioeconomic Characteristics and School-Level Performance Assessment Results

Thomas R. Guskey

University of Kentucky

Send correspondence to

Thomas R. Guskey
College of Education
University of Kentucky
Lexington, KY 40506
Phone: 606/257-8666
E-mail: GUSKEY@POP.UKY.EDU

This paper was presented at the annual meeting of the American Educational Research Association, Chicago, IL, March 1997.

TM026501

BEST COPY AVAILABLE

Abstract

The purpose of this study was to investigate the relationship between selected contextual and socioeconomic variables and school-level results from a statewide, performance-based student assessment system. Multiple regression analyses were conducted using data gathered over a three year period from 49 schools in one school district. Results showed that a single socioeconomic variable, the percent of students qualifying for free or reduced lunch benefits, explained a large portion of the variation in scores at all school levels. Contextual or socioeconomic indicators were not, however, predictive of improvements in scores from year to year. Possible explanations for these results and school level differences are discussed, along with implications for performance-based student assessment systems.

The Relationship Between Socioeconomic Characteristics and School-Level Performance Assessment Results

Performance assessment is a dominant force in education today. Spurred by evidence of the serious limitations and numerous unintended, negative consequences of standardized, multiple-choice tests, performance assessments are advocated as far more "authentic" measures of student learning (Meisels, Dorfman, & Steele, 1995; Wiggins, 1993). When compared to traditional tests, performance assessments appear less stigmatizing, more instructionally relevant, more adaptable to individual student differences, more faithful to the richness and complexity of real-world problem-solving, and more reflective of the actual quality of student understanding. As Bond (1995, p. 21) points out, however, these qualities are more "articles of faith" than demonstrated characteristics. Little evidence exists currently that performance assessments are truly less stigmatizing or that they are a more accurate reflection of student learning.

Another suggested advantage of performance assessments is that they are, at least potentially, less biased and more fair to traditionally disadvantaged students. This is because performance assessments, when properly used, can merge instruction and assessment, rather than simply test abilities (such as recall of decontextualized information) that are only remotely connected to the everyday experiences of most students (Simmons & Resnick, 1993). Although it would be premature to say that any definitive conclusion has been reached regarding this claim, the information gathered to date is not particularly promising. Jaeger (1992), for example, reports a rather consistent pattern of findings linking poverty rate and other family structure variables to student performance on problem solving tasks in international mathematics assessments. Similarly,

analyses of data from the National Assessment of Educational Progress show that after correcting for reliability, differences in scores on extended-response essays between African American students and White students actually exceed those on the multiple-choice reading assessment (Baker, O'Neil, & Linn, 1991; Linn, Baker, & Dunbar, 1991).

This study was designed to provide additional insights into this complex interaction. Specifically, it's purpose was to investigate the relationship between selected contextual and socioeconomic variables, and school-level results from a statewide performance-based assessment system. We sought to determine the degree to which school-level performance assessment results could be predicted from data on selected contextual factors and the socioeconomic characteristics of a school's student population. A previous study by Guskey and Kifer (1990) showed that certain socioeconomic indices could be used to explain a large portion of the variation (approximately 42%) among school districts in the results from a basic skills testing program used formerly in the same state. The developers of the current performance-based assessment system hoped that the change in assessment content, format, and procedures might alter this trend.

The Assessment

A major element of the Kentucky Education Reform Act is a statewide student assessment program known as the Kentucky Instructional Results Information System, or KIRIS. This multi-faceted performance assessment system has three components (see Guskey, 1994a). The first is evaluations of portfolios of students' work in the areas of writing, mathematics, science, and social studies. The second is based on students' scores from "on-demand assessments" in the areas of reading, mathematics, science, social studies, arts and humanities, practical living and vocational

skills. These assessments include both multiple-choice and open-ended items similar to those used in the National Assessment of Educational Progress. The third component involves students' performance on a series of "performance events" in the areas of mathematics, science, social studies, arts and humanities, practical living and vocational skills.

Results from these three components of KIRIS are combined with a series of "non-cognitive indicators" (e.g., school attendance, retention, dropout rate, etc.) to calculate an "accountability index score" for every public school in the Commonwealth. In determining the "accountability index" in the earliest versions of KIRIS, performance assessment results contribute 5/6th while the "non-cognitive indicators" contribute 1/6th.

Scores on the "accountability index" can range from 0 to 100, and are based on the goal of having *all* students (100%) score at the "proficient" level on the assessments within a period of 20 years (Kifer, 1994). In practice, however, not *all* students need to score at the "proficient" level for a school to reach an accountability index score of 100. Students who score at the highest "distinguished" level on the assessments balance the scores of those students who score at the lower levels of "novice" and "apprentice."

The first year the KIRIS assessments were administered in 1992, results were used to determine a "baseline" accountability index score for each school. From this baseline, all schools were expected to show improvement. The amount of improvement expected was determined by calculating a "threshold" score for each school. Thresholds were established by first computing the gain a school would have to make to reach an accountability index score of 100 within a time period of 20 years. This total amount is then divided into ten, equal increments. Adding one increment to a school's "baseline" yields that school's "threshold" for the next biennial assessment. So, for example, if a school's "baseline" accountability index score in 1992 was 30, it's "threshold"

for 1994 would be 37 ([100-30]/10 = 7; 30 + 7 = 37). Based on a school's results, new thresholds are then recalculated for the next assessment cycle.

To avoid the problems associated with year to year score fluctuations (Mandeville & Khosrow, 1988), scores from two years are combined (averaged) to obtain each school's new accountability index score. This score is then compared to the school's "threshold" to determine if expected progress has been made. In other words, the accountability associated with KIRIS is not based on *achievement* per se; rather, it is based on *improvement*. Each school is compared to itself and is expected to show consistent improvement over time.

The reform legislation further specifies that results from the assessments will be used to grant financial rewards to schools that show significant improvements (i.e., make gains beyond their "threshold") and to levy sanctions against those that fail to show progress. In other words, the assessment program is also "high stakes" (Trimble, 1994).

Data Source

The principal source of data for this study was the 1993, 1994, and 1995 KIRIS results from one of the largest school districts in the Commonwealth. This county school district includes a metropolitan area and its surrounding suburbs. Nearly 34,000 students are enrolled in the district's 33 elementary schools (grades K-5), 11 middle schools (grades 6-8), and 5 high schools (grades 9-12). The performance data used in this study include the accountability index scores for each of the 49 schools in the district for each of the three years. Simple gain scores also were calculated by subtracting the previous years' score from each school's accountability index score.

The reform legislation originally specified that assessment results from grades 4, 8, and 12 would be used to calculate the accountability index scores for elementary, middle, and high schools,

respectively. But several changes have been made to this policy since the original legislation was passed. Specifically, the mathematics portfolio assessment has been moved from grade 4 to grade 5, and the performance assessments and on-demand assessments for high school students have been moved from grade 12 to grade 11. Furthermore, during the years for which these data were gathered, only student portfolios in the areas of writing and mathematics, and on-demand assessments in the areas of reading, mathematics, science, and social studies, were considered in calculating accountability index scores.

The contextual and socioeconomic data regarding the student population of each school were obtained from district records. These data include: school enrollment, percent of minority students, percent of students qualifying for free or reduced lunch benefits, percent of students enrolled in special education, percent mobility (students who withdraw or are added to a school's enrollment in a given year), percent retention (students not promoted to the next grade[1]), suspension days (number of days students are suspended), expulsion days (number of days students are expelled), and pupil/teacher ratio. Inspection of district records showed that for individual schools, these school-based measures varied little from 1993 to 1995. Therefore, the most recent data (1995) from each school were used in all analyses.

Method

The data were analyzed in four stages. First, means and standard deviations were computed for all measures to determine the extent of variation among the schools. Because of differences in the procedures used to calculate the accountability index scores for elementary, middle, and high schools, data from each of these levels were analyzed separately at all stages in the analysis.

Second, correlation coefficients were computed between all measures to determine the degree of linear relationship. Third, multiple regression analyses were conducted in which the KIRIS accountability index scores from each year were regressed on the entire group of contextual and socioeconomic indices. Separate multiple regression analyses also were conducted using gain scores as the dependent variables.

Finally, contextual and socioeconomic indices were eliminated from each regression model in stepwise fashion to produce the most parsimonious models for predicting accountability index scores and gain scores.

Results

Descriptive statistics are illustrated in Table 1 and show substantial variation among schools at all levels on nearly every variable. Exceptions include the expulsion days at the elementary level (0 for all schools) and pupil/teacher ratio (nearly the same for schools at each level). On most measures, variation is greatest among the elementary schools and least among the high schools. Enrollment, for example, ranges from 222 to 884 for elementary schools, from 245 to 990 for middle schools, and from 1277 to 1606 for high schools. Similarly, the percent of minority students varies from 5% to 72% for elementary schools, from 12% to 45% for middle schools, and from 16% to 41% for high schools. Accountability index scores show considerable variation as well, with 1995 scores ranging from 29.2 to 72.8 for elementary schools, from 38.3 to 75.6 for middle schools, and from 36.2 to 56.3 for high schools.

[Insert Table 1]

The results of the correlation analyses are shown in Tables 2-4. These analyses were somewhat restricted at the high school level because there are only five high schools in this district. Nevertheless, several interesting patterns emerged.

As expected, many of the contextual and socioeconomic indicators are highly correlated at all three levels of schools. Particularly interesting is the relationship between the primary indicator of poverty and that of minority enrollment. Correlations between percent of students in a school qualifying for free or reduced lunch benefits and percent of minority students are .82, .92, and .96 for elementary, middle, and high schools, respectively. Also highly correlated at most levels with percent of students qualifying for free or reduced lunch benefits are percent of students in special education programs, mobility rates, and retention rates. In other words, in this school district, schools with high percentages of students from economically disadvantaged backgrounds also tend to have more minority students, more students placed in special education programs, more mobility among students, and more students retained.

Also as expected, accountability index scores from the three years are highly correlated. This shows that schools with high accountability scores tend to maintain relatively high scores while those with low scores remained relatively low. Gain scores, on the other hand, were uncorrelated with accountability index scores and negatively correlated across years. In other words, schools' scores on the accountability index had no relation to the gain they were able to achieve from one year to next. Unexpectedly, however, the gains schools achieved from 1993 to 1994 on the accountability index were negatively related to gains they achieved from 1994 to 1995 (elementary: $r = -.54$; middle: $r = -.40$; high: $r = -.47$). This means that gains from year to year are not simply unstable; they are *inversely related*. In general, the more a school gained one year, the less it gained the following year, and vice versa.

[Insert Tables 2, 3, & 4]

Among elementary schools, the correlations between accountability index scores and enrollment were positive but not statistically significant (r = +.26 to +.33). Correlations between other contextual and socioeconomic variables and accountability index scores, however, were generally negative and large in magnitude (r = -.53 to -.80). This means, for example, that as the percent of students qualifying for free and reduced lunch benefits in a school increased, accountability index scores generally decreased. These trends changed only slightly for middle schools, where again the correlations between socioeconomic indices and accountability index scores were negative and equally large in magnitude (r = -.60 to -.96). The contextual variables of mobility, retention, and suspensions also were negatively correlated with accountability index scores at this level (r = -.64 to -.91).

At the high school level, correlations between contextual and socioeconomic variables and accountability index scores were similarly negative and large in magnitude (r= -.66 to -.97). A more detailed analysis of the high school data showed, however, that four of the five schools clustered closely on nearly all measures while one school was a consistent outlier. This pattern, combined with the small number of high schools included in the analysis, may account for the unusually high correlations between several variables.

Interestingly, correlations between the contextual and socioeconomic variables and gain scores were consistently small and not statistically significant at all school levels. In other words, improvement in accountability index scores appears to be independent of the contextual and socioeconomic characteristics of a school's student population.

Regression analyses were then conducted, first using the entire set of contextual and socioeconomic measures as independent variables, and accountability index scores or gain scores as the dependent variable. Predictor variables next were eliminated in stepwise fashion, testing the reduction in the R-square value. It was discovered that a single socioeconomic variable, percent of students qualifying for free or reduced lunch benefits, provided the most parsimonious model for predicting accountability index scores. This is undoubtedly due to the high degree of multi-collinearity among the contextual and socioeconomic indicators included in these data.

Results from the regression analyses showed the percent variance explained (R-squared) in accountability index scores for the three years ranged from 51% to 66% for elementary schools, from 60% to 81% for middle schools, and from 61% to 78% for high schools. In all cases, these values are higher than the R-squared value (42%) obtained in the previously noted research that involved a statewide, multiple-choice, basic skills testing program (Guskey & Kifer, 1990). Thus the change in assessment format appears *not* to have altered the relationship between this socioeconomic indicator and assessment results.

It should also be noted, however, for elementary and middle schools the R-square values were highest for 1993 results and lowest for 1995 results. At the high school level, the opposite trend was noted: the R-square value was lowest for 1993 results and highest for 1995 results. Although the differences among these R-square values are not statistically significant, this finding provides preliminary evidence that over time, this socioeconomic indicator may be becoming a less powerful predictor of accountability index scores for elementary and middle schools. Whether the opposite trend for high schools is generalizable or unique to the high schools in this particular district is unknown.

Regression analyses with gain scores yielded much smaller and nonsignificant R-square values at all school levels. This shows that socioeconomic indicators are not powerful predictors of the gains schools can achieve on these performance-based student assessments. Results from analyses of the most current accountability index scores (1995) are illustrated in Table 5.

[Insert Table 5]

These mixed findings are surprising but generally optimistic. They indicate that like the results from the basic skills testing program used formerly in the Commonwealth, socioeconomic indicators are powerful predictors of a school's initial achievement standing. Such indicators, however, do not predict the progress or improvement a school can make. The contextual and socioeconomic variables considered in this study were unrelated the improvements schools made, measured in terms of results on performance-based student assessments. The negative relationship between the gains made by schools in two consecutive years, however, remains puzzling.

Conclusions and Implications

The results from this investigation show that although socioeconomic indicators explain much of the variation in elementary, middle, and high schools' initial level of achievement in a high-stakes, performance-based student assessment program, they do not explain the variation in score improvements made by schools in this school district. In other words, the contextual characteristics of a school and the socioeconomic characteristics of the school's student population are unrelated to the magnitude of improvement attained. This is a very positive finding from the perspective of advocates of performance-based assessment systems who have hoped this change in assessment

content, format, and procedures would result in fair and instructionally valid measures of student learning. These results also suggest that this high-stakes, performance-based assessment system may be prompting schools at all levels to make changes that lead to improvements in the performance of their students.

At the same time, there are two aspects of these results that deserve special attention. First is the difference in trends for elementary and middle schools versus high schools. Among the elementary and middle schools, the relationship between socioeconomic indicators and student performance results is smaller in magnitude and appears to be diminishing. Among the high schools, however, this relationship is larger in magnitude and appears to be increasing. This finding is particularly troubling if one considers the probable impact of school dropouts on the assessment results.

Approximately 22 percent of the high school students enrolled in the ninth grade in this school district fail to graduate within four years. Those students who dropout or otherwise fail to graduate typically have low levels of achievement and poor academic records. They also include disproportionately large numbers of economically disadvantaged and minority students. The effect of dropouts, therefore, is to reduce the variance in both achievement measures and socioeconomic indicators. This reduction in variance, in turn, should reduce the magnitude of the correlation between these measures.

In this particular instance, however, the relationship between socioeconomic indicators and achievement measures is *greater* in magnitude at the high school level than at the elementary and middle school levels, and appears to be *increasing*. The nature of these data, particularly the divergence of one of the high schools from the other four on many of the measures, confounds efforts to provide precise explanations. Nevertheless, this finding clearly calls for an investigation

in this district of possible unintended consequences stemming from formal or informal tracking practices, or other inequitable opportunities for students to learn (Winking & Bond, 1996).

The second aspect of these results deserving special attention is the negative correlation between accountability index score gains from one year to the next for all school levels. In other words, the gains schools achieved from 1993 to 1994 on the accountability index are negatively related to gains achieved from 1994 to 1995. Some have suggested this negative correlation is an indication that the gains are spurious or artificial. Others argue it may be due to differences in the quality or difficulty of certain aspects of the assessment system from year to year (Kentucky Institute for Education Reform, 1995). A closer examination of school-by-school results yields still another plausible explanation.

The Kentucky Instructional Results Information System, with its accompanying accountability system, is based on the assumption "What you test is what you get." That is, in a high stakes environment where rewards or sanctions are tied to test performance, teachers will alter their instruction to match the content and format of the tests. Although this is generally true with regard to basic skills testing, current evidence indicates it may not be true when more authentic, performance-based assessments are used (Linn, 1993). Most teachers have scant knowledge, personal background, experience, or formal training in performance assessments, or in how to use them as instructional tools (Rothman, 1995). In addition, the majority of teachers indicate they do not have sufficient time to administer and score such assessments, nor do they have access to appropriate and well-aligned instructional materials (Vitali, 1993).

To make these substantial changes in their instructional practices, most teachers need extended time and sustained, high quality professional development (Guskey, 1994b). But instead of providing such time and resources, pressure for immediate improvement in scores has prompted

many schools in Kentucky to develop professional development programs that focus narrowly on the particular assessment formats and scoring procedures included in KIRIS (Oldham, 1994). A recent study showed, for example, that all surveyed principals reported encouraging teachers to use special test-preparation materials and to teach explicit test-taking skills (Koretz, Barron, Mitchel, & Stecher, 1996). Consequently, teachers teach their students precise strategies for tailoring responses to specific scoring rubrics. Although this generally leads to improved scores, such improvements are typically modest and short-lived (Guskey & Oldham, 1997).

In examining school-by-school data, it appears about one-third of the schools in this district may have improved their assessment results in the first year by teaching students these specific response strategies. As a result, they experienced a first year increase in their accountability index scores. In the second year, however, improvements leveled off. Another third of the schools may not have provided this focus until the second year. Their results were relatively flat during the first year but saw an increase in year two. The final third of the schools apparently did not provide this focus or any other meaningful alternative to improve results. Hence, they saw little overall gain.

The gains achieved by schools in one year were not, therefore, followed by a loss the next year, or vice versa. Rather, it is simply that among the schools in this school district, significant gains are typically followed or preceded by a period of relatively flat improvement. Perhaps this is because significant gains are difficult to sustain from year to year. Perhaps steady improvements require more time and resources than have been provided. Or perhaps, as suggested, the gains evidenced thus far have been attained simply by teaching students specific response strategies, and the hoped-for improvements in teaching practices that will lead to greater depth in students' understanding have yet to be realized. Additional studies exploring these issues are clearly needed,

especially efforts that might include the disaggregation of accountability indices to consider changes in results from year to year in specific subject areas.

Because the results from this study are based on data gathered from a single school district, generalizations are limited. Nevertheless, it does provide preliminary evidence on the relationship between contextual variables and socioeconomic indicators, and school-level results from a performance-based student assessment program. Although performance assessments clearly do not provide a solution to all of the problems schools face in their efforts to improve the quality of education provided students, they may be moving us in the right direction.

Endnotes

[1] The percent retention for elementary schools is based *only* on students retained in grades 4 and 5. From kindergarten through grade 3 students are enrolled in a "non-graded primary school program," which is mandated as part of the Kentucky Education Reform Act. Under this program, students who are "given another year" in the primary grades are not counted in determining a school's retention rate.

References

Baker, E. L., O'Neil, H. F., & Linn, R. L. (1991, April). *Policy and validity prospects for performance-based assessment*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Bond, L. (1995). Unintended consequences of performance assessment: Issues of bias and fairness. *Educational Measurement: Issues and Practice*, *14*(4), 21-24.

Guskey, T. R. (Ed.) (1994a). *High stakes performance assessment: Perspectives on Kentucky's educational reform*. Thousand Oaks, CA: Corwin Press.

Guskey, T. R. (1994b). What you assess may not be what you get. *Educational Leadership*, *51*(6), 51-54.

Guskey, T. R., & Kifer, E. W. (1990). Ranking school districts on the basis of statewide test results: Is it meaningful or misleading? *Educational Measurement: Issues and Practice*, *9*(1), 11-16.

Guskey, T. R., & Oldham, B. R. (1997, in press). Despite the best intentions: Inconsistencies among components in Kentucky's systemic reform. *Educational Policy*.

Jaeger, R. M (1992). World class standards, choice, and privatization: Weak measurement serving presumptive policy. *Phi Delta Kappan*, *74*(2), 118-128.

Kentucky Institute for Education Reform (1995, January). *An independent evaluation of the Kentucky Instructional Results Information System conducted by the Evaluation Center, Western Michigan University*. Frankfort, KY: The Kentucky Institute for Education Reform.

Kifer, E. (1994). Development of the Kentucky Instructional Results Information System (KIRIS). In T. R. Guskey (Ed.), *High stakes performance assessment: Perspectives on Kentucky's educational reform* (pp. 7-18). Thousand Oaks, CA: Corwin Press.

Koretz, D. M., Barron, S., Mitchell, K. J., & Stecher, B. M. (1996). *Perceived effects of the Kentucky Instructional Results Information System (KIRIS)*. Santa Monica, CA: Rand.

Linn, R. L. (1993). Educational assessment: Expanded expectations and challenges. *Educational Evaluation and Policy Analysis, 15*, 1-16.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.

Mandeville, G. K., & Khosrow, H. (1988, April). *Measuring school effectiveness using hierarchical linear models*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans.

Meisels, S. J., Dorfman, A., & Steele, D. (1995). Equity and excellence in group-administered and performance-based assessments. In M. T. Nettles & A. L. Nettles (Eds.), *Equity and excellence in educational testing and assessment* (pp. 243-261). Boston: Kluwer.

Oldham, B. R. (1994). A school district's perspective. In T. R. Guskey (Ed.), *High stakes performance assessment: Perspectives on Kentucky's educational reform* (pp. 55-63). Thousand Oaks, CA: Corwin Press.

Rothman, R. (1995). *Measuring up: Standards, assessment, and school reform*. San Francisco: Jossey-Bass.

Simmons, W., & Resnick, D. (1993, March). *National standards, assessment, and equity*. Paper presented at the Symposium on Equity and Educational Testing and Assessment. The Ford Foundation: Washington, DC.

Trimble, C. S. (1994). Ensuring educational accountability. In T. R. Guskey (Ed.), *High stakes performance assessment: Perspectives on Kentucky's educational reform* (pp. 37-54). Thousand Oaks, CA: Corwin Press.

Vitali, G. J. (1993). *Factors influencing teachers' assessment and instructional practices in an assessment-driven educational reform*. Doctoral dissertation, University of Kentucky.

Wiggins, G. (1993). Assessment: Authenticity, context, and validity. *Phi Delta Kappan, 75*(3), 200-214.

Winking, D. L., & Bond, L. A. (1996). *Transforming teaching and learning in urban schools through alternative assessment*. Oak Brook, IL: North Central Regional Educational Laboratory.

Table 1

Descriptive Statistics for Elementary, Middle, and Secondary Schools

Variable	Elementary (n = 33)			Middle (n = 11)			Secondary (n = 5)		
	X	sd	range	X	sd	range	X	sd	range
Enrollment	505.1	190.3	222 - 884	708.3	207.6	245 - 990	1503.4	136.5	1277 - 1606
% Minority	29.5	17.5	4.9 - 71.7	26.1	12.6	11.8 - 44.7	26.4	10.1	16.0 - 41.2
% Free Lunch	44.0	26.5	4.0 - 94.0	30.9	19.3	5.0 - 64.0	21.2	9.9	12.0 - 38.0
% Special Ed.	13.3	6.1	3.0 - 30.0	10.4	5.4	1.0 - 17.0	8.6	1.7	7.0 - 11.0
% Mobility	2.5	1.8	1.0 - 7.0	10.4	6.0	0.8 - 22.8	8.0	1.9	6.0 - 10.3
% Retained	0.8	2.2	0.0 - 11.8	3.4	2.5	0.0 - 7.5	8.6	1.9	6.3 - 19.5
Suspended	6.2	9.0	0.0 - 45.0	202.5	184.5	5.0 -508.0	620.0	519.0	223.0 - 1530.0
Expelled	0.0	0.0	0.0 - 0.0	0.5	0.8	0.0 - 2.0	2.6	2.5	1.0 - 7.0
Pupil/Teacher	16.9	2.0	12.6 - 20.9	15.0	1.4	13.1 - 16.9	18.4	.8	17.5 - 19.7
93 Index	36.1	8.7	25.8 - 60.8	43.8	13.6	27.1 - 69.8	39.5	5.3	31.3 - 44.5
94 Index	41.9	9.7	24.6 - 62.1	49.0	11.9	34.4 - 75.9	48.4	4.0	42.4 - 52.8
95 Index	47.1	8.9	29.2 - 72.8	50.6	12.5	38.3 - 75.6	49.6	8.1	36.2 - 56.3
93-94 Gain	5.8	5.8	-4.0 - 18.3	5.2	3.5	-0.7 - 10.2	8.9	5.0	3.5 - 14.3
94-95 Gain	5.2	4.4	-4.2 - 15.3	1.6	4.4	-3.5 - 10.7	1.2	5.2	-6.0 - 8.3
93-95 Gain	11.0	5.1	2.5 - 20.3	6.8	5.5	-1.0 - 16.2	10.1	5.3	4.2 - 15.6

Table 2

Correlation Coefficients Among Elementary School Measures
(n = 33)

Variable	Enroll.	% Min.	% Free	% Spec.	% Mob.	% Ret.	Susp.	Pup/Tch	93 Index	94 Index	95 Index	93-94 Gain	94-95 Gain
% Minority	-.58**												
% Free Lunch	-.62**	.82**											
% Special Ed.	-.58**	.64**	.86**										
% Mobility	.19	-.36*	-.40*	-.21									
% Retained	-.34	.22	.26	.32	-.15								
Suspended	.01	-.03	.13	.01	-.37*	-.15							
Pupil/Teacher	.84**	-.56**	-.70**	-.64**	.25	-.23	-.07						
93 Index	.32	-.64**	-.80**	-.67**	.38*	-.25	-.19	.42*					
94 Index	.33	-.59**	-.77**	-.72**	.38*	-.13	-.30	.46**	.80**				
95 Index	.26	-.53**	-.71**	-.65**	.31	-.18	-.29	.42*	.83**	.89**			
93-94 Gain	.07	-.02	-.09	-.20	.06	.17	-.22	.14	-.16	.46**	.23		
94-95 Gain	-.20	.21	.25	.27	-.20	-.09	.08	-.16	-.08	-.40*	.06	-.54**	
93-95 Gain	-.10	.17	.12	.00	-.10	.12	-.18	.03	-.25	.18	.32	.68**	.25

* = $p < .05$
** = $p < .01$

Table 3

Correlation Coefficients Among Middle School Measures
(n = 11)

Variable	Enroll.	% Min.	% Free	% Spec.	% Mob.	% Ret.	Susp.	Exp.	Pup/Tch	93 Index	94 Index	95 Index	93-94 Gain	94-95 Gain
% Minority	-.22													
% Free Lunch	-.04	.92**												
% Special Ed.	.09	.81**	.91**											
% Mobility	.08	.78**	.92**	.89**										
% Retained	.12	.79**	.85**	.94**	.81**									
Suspended	.12	.61*	.68*	.80**	.57	.75**								
Expelled	.06	.24	.48	.47	.63*	.32	.42							
Pupil/Teacher	.59	-.49	-.55	-.48	-.50	-.48	-.37	-.24						
93 Index	-.12	-.78**	-.90**	-.96**	-.91**	-.90**	-.68*	-.42	.53					
94 Index	-.18	-.64*	-.82**	-.92**	-.89**	-.86**	-.65*	-.43	.52	.97**				
95 Index	-.36	-.60*	-.77**	-.88**	-.85**	-.77**	-.64*	-.42	.25	.91**	.94**			
93-94 Gain	-.12	.80**	.69*	.57	.48	.52	.42	.16	-.30	-.56	-.34	-.34		
94-95 Gain	-.54	.03	.03	.03	.03	.16	-.06	-.04	-.71*	-.05	-.06	.29	-.40	
93-95 Gain	-.50	.54	.47	.39	.33	.46	.22	.07	-.75**	-.39	-.27	.01	.61*	.77**

* = $p < .05$
** = $p < .01$

Table 4

Correlation Coefficients Among High School Measures
(n = 5)

Variable	Enroll.	% Min.	% Free	% Spec.	% Mob.	% Ret.	Susp.	Exp.	Pup/Tch	93 Index	94 Index	95 Index	93-94 Gain	94-95 Gain
% Minority	-.79													
% Free Lunch	-.90•	.96••												
% Special Ed.	-.92•	.76	.82											
% Mobility	-.72	.35	.59	.63										
% Retained	-.68	.50	.70	.54	.93•									
Suspended	.30	.04	.04	-.40	.08	.37								
Expelled	-.94•	.90•	.97••	.79	.62	.70	-.04							
Pupil/Teacher	.12	-.20	-.31	-.13	-.61	-.76	-.76	-.20						
93 Index	.66	-.77	-.78	-.39	-.25	-.44	-.10	-.86	-.01					
94 Index	.94•	-.73	-.80	-.99••	-.62	-.50	-.50	-.80	.01	.44				
95 Index	.97••	-.74	-.88•	-.80	-.73	-.72	-.20	-.96••	.13	.76	.84			
93-94 Gain	.05	.22	.18	-.37	-.22	.06	.50	.26	.02	-.70	.33	-.13		
94-95 Gain	.79	-.60	-.76	-.48	-.66	-.74	-.06	-.87	.20	.86	.54	.91•	-.47	
93-95 Gain	.82	-.40	-.56	-.83	-.86	-.66	.42	-.60	.21	.16	.85	.76	.50	.53

• = $p < .05$
•• = $p < .01$

Table 5

Regression Analyses Results for 1995 Accountability Index Scores

Source	df	SS	MS	F	p	R^2
Elementary Schools						
% Free/Reduced	1	1286.9	1286.9	31.64	0.000	50.5
Error	31	1261.0	40.7			
Total	32	1548.0				
Middle Schools						
% Free/Reduced	1	924.2	924.2	13.26	0.005	59.6
Error	9	627.5	69.7			
Total	10	1551.7				
High Schools						
% Free/Reduced	1	203.4	203.4	10.66	0.047	78.0
Error	3	57.3	19.1			
Total	4	260.7				